ISO 19115-1

ISO Standards are eligible for systematic revisions every five years. The revision process can be initiated by a single member voting in favor of a revision. The Principle ISO Metadata Standard (ISO 19115) was revised over the last several years and the result (ISO 19115-1) will soon be an new international standard replacing ISO 19115. NASA and NOAA participated in the revison process which resulted in consideration of their metadata requirements in the process. Many of the revisions included in ISO 19115-1 were suggested by NASA and NOAA. Some of the successes are described here.

Keyword Types

Keywords from shared vocabularies are critical for consistent data discovery interfaces. The NASA Global Change Master Directory has developed shared vocabularies that include consistent values for many important types of keywords. These vocabularies are used extensively throughout the GCMD community and are included in GCMD's Directory Interchange Format (DIF).

ISO 19115 includes a shared vocabulary (codelist) for keyword types that includes five values: discipline, place, stratum, temporal, and theme. This codelist was extended in ISO 19115-1 to include the GCMD keyword types and others: dataCentre, featureType, instrument, platform, project, service, product, subTopicCategory, taxon. This will allow direct translation of keywords from the GCMD metadata into ISO, greatly simplifying the translation and, more importantly, maintaining a consistent conceptual interface for those users as they begin to use ISO.

Multiple names for measured or calculated parameters

Names of measured and calculated parameters in NASA datasets are chosen by science teams in order to fit in with their systems and approaches. In most cases these names are not consistent with standard names from shared vocabularies. In order to be compliant with those vocabularies, these parameters need multiple names or aliases. ISO 19115 supports only one name for each parameter in a dataset. This is the local name for the parameter. ISO 19115-1 adds the ability to attach any number of identifiers to parameters. This enables interoperability across multiple shared vocabularies and systems that use them.

Using NASA identifiers

The ISO Metadata Standards include identifiers for a large number if components. In 19115 these identifiers included a code and a citation for the authority of the code. This worked well in some cases, but it was difficult to simply associate a namespace with a code. In the revision the identifier has been expanded to include a namespace (called codespace), a version, and a description. This makes the identifiers easier to use in a web environment as well as making them easier for users to understand. Many NASA identifiers include short and long names. In the translation to ISO the short name becomes the code and the long name becomes the description.

Multiple metadata dates

Metadata for NASA data and products is created, revised, and shared using several different systems.Tracking of metadata history and migration between Data Centers and Repositories is important for ensuring that users get current information. ISO 19115 includes the capability to define several types of dates and time ranges for resources, but only includes a creation date for the metadata. The creation date is important, but having the flexibility to define other dates types is also useful. ISO 19115-1 includes seventeen date types (creation, publication, lastUpdate, nextUpdate, superseded, …) and any of them can be associated with the metadata.

References to documentation or other information that is published in the literature

Information required for using and understanding NASA data and products comes in many forms. Structured and standard metadata can only include a portion of this documentation. Other important information might published in user guides, data dictionaries or scientific papers. ISO 19115-1 introduces the capability to add citations to those materials to the metadata record which becomes a discoverable source for links to all related documentation.

Data usage and limitations identified by users

NASA data are used in a variety of sometimes unexpected ways by diverse users all over the world. In some cases, those users identify limitations or usability problems. ISO 19115 includes information about when and how users are using data as well as limitations that they might identify as they use it. ISO 19115-1 introduces the capability for data providers to include responses to those limitations when they have been addressed. This is an important mechanism for helping users understand how datasets are improved or updated in response to their input.

Special constraints or use limitations for different time periods

Some datasets have special constraints for selected time periods. For example, some data can only be used by the principle investigator or members of the science team for some period of time after they are collected or their may be instrument problems that limit the utility of the data during some time period. Use of other data sets might be limited by license arrangements. ISO 19115-1 introduces the capability to specify temporal and spatial extents of constraints. It also allows the specification of specific constraints for particular users.

Do you need to describe coverages with physical measurements, quality information, auxiliary information, reference information, or model results?

Many NASA data and products are provided in grids and accompanied by gridded quality information or auxiliary data. ISO 19115 allows the description of multiple electromagnetic bands in coverages which can be useful for instrument data but difficult to use when datasets include coverages with multiple layers with different types of information. ISO 19115-1 extends the MD_Band object to include this capability. It includes simple descriptive statistics for each layer and allows inclusion of custom layer attributes as well.